

# Particle size in aerosols

The particle size distribution within aerosols is modelled using the log-normal distribution. It is reasonable to assume that the diameters of particles identically distributed and independent. Using this model, find the joint probability density function of particle sizes in a sample of  $n$  particles, and state the parameters of the model.

Recall that the p.d.f. of the log-normal distribution is

$$f_Y(y) = \begin{cases} \frac{1}{(2\pi\sigma^2)^{1/2}} \frac{1}{y} \exp\left(-\frac{(\log y - \mu)^2}{2\sigma^2}\right) & \text{if } y \in (0, \infty) \\ 0 & \text{otherwise.} \end{cases}$$

The parameters of this distribution, and hence also the parameters of our model, are  $\mu \in \mathbb{R}$  and  $\sigma \in (0, \infty)$ .

Since particle sizes are independent, the joint p.d.f. of

$$\mathbf{Y} = (Y_1, Y_2, \dots, Y_n),$$

where  $Y_i$  is the diameter of the  $i^{th}$  particle, is

$$\begin{aligned} f_{\mathbf{Y}}(y_1, \dots, y_n) &= \prod_{i=1}^n f_{Y_i}(y_i) \\ &= \begin{cases} \frac{1}{(2\pi\sigma^2)^{n/2}} \frac{1}{y_1 y_2 \dots y_n} \exp\left(-\sum_{i=1}^n \frac{(\log y_i - \mu)^2}{2\sigma^2}\right) & \text{if } y_i > 0 \text{ for all } i \\ 0 & \text{otherwise.} \end{cases} \end{aligned}$$

# Radioactive decay

The time taken for a single atom of, say, carbon-15 to decay can be modelled as an exponential random variable, with unknown parameter  $\lambda \in (0, \infty)$ . The parameter  $\lambda$  is known as the 'decay rate'. The times at which atoms decay are known to be independent.

Using this model, find the likelihood function for the decay times of a sample of  $n$  carbon-15 atoms.

# Radioactive decay

The time taken for a single atom of, say, carbon-15 to decay can be modelled as an exponential random variable, with unknown parameter  $\lambda \in (0, \infty)$ . The parameter  $\lambda$  is known as the 'decay rate'. The times at which atoms decay are known to be independent.

Using this model, find the likelihood function for the decay times of a sample of  $n$  carbon-15 atoms.

The decay time  $X_i$  of the  $i^{th}$  atom is exponential with parameter  $\lambda \in (0, \infty)$ , and therefore has p.d.f.

$$f_{X_i}(x_i; \lambda) = \begin{cases} \lambda e^{-\lambda x_i} & \text{if } x_i > 0 \\ 0 & \text{otherwise.} \end{cases}$$

Since each atom decays independently, the joint distribution of  $\mathbf{X} = (X_i)_{i=1}^n$  is

$$\begin{aligned} f_{\mathbf{X}}(\mathbf{x}; \lambda) &= \prod_{i=1}^n f_{X_i}(x_i; \lambda) = \begin{cases} \prod_{i=1}^n \lambda e^{-\lambda x_i} & \text{if } x_i > 0 \text{ for all } i \\ 0 & \text{otherwise.} \end{cases} \\ &= \begin{cases} \lambda^n \exp(-\lambda \sum_{i=1}^n x_i) & \text{if } x_i > 0 \text{ for all } i \\ 0 & \text{otherwise.} \end{cases} \end{aligned}$$

Since each atom decays independently, the joint distribution of  $\mathbf{X} = (X_i)_{i=1}^n$  is

$$\begin{aligned} f_{\mathbf{X}}(\mathbf{x}; \lambda) &= \prod_{i=1}^n f_{X_i}(x_i; \lambda) = \begin{cases} \prod_{i=1}^n \lambda e^{-\lambda x_i} & \text{if } x_i > 0 \text{ for all } i \\ 0 & \text{otherwise.} \end{cases} \\ &= \begin{cases} \lambda^n \exp(-\lambda \sum_{i=1}^n x_i) & \text{if } x_i > 0 \text{ for all } i \\ 0 & \text{otherwise.} \end{cases} \end{aligned}$$

Therefore, the likelihood function is

$$L(\lambda; \mathbf{x}) = \begin{cases} \lambda^n \exp(-\lambda \sum_{i=1}^n x_i) & \text{if } x_i > 0 \text{ for all } i \\ 0 & \text{otherwise.} \end{cases}$$

The range of possible values for  $\lambda$  is  $\Theta = (0, \infty)$ .

Suppose that we have sampled the decay times of 15 carbon-15 atoms (in seconds, given to two decimal places), and found them to be

$$\mathbf{x} = \{0.50, 2.19, 0.88, 4.06, 9.75, 2.62, 0.13, 2.70, 0.03, 0.28, 4.15, 9.52, 2.67, 3.79, 4.31\}.$$

Find the maximum likelihood estimator of  $\lambda$ , based on this data.

Suppose that we have sampled the decay times of 15 carbon-15 atoms (in seconds, given to two decimal places), and found them to be

$$\mathbf{x} = \{0.50, 2.19, 0.88, 4.06, 9.75, 2.62, 0.13, 2.70, 0.03, 0.28, 4.15, 9.52, 2.67, 3.79, 4.31\}.$$

Find the maximum likelihood estimator of  $\lambda$ , based on this data.

For this data,  $\sum_{i=1}^{15} x_i = 47.58$ , so our likelihood function is

$$L(\lambda; \mathbf{x}) = \lambda^{15} e^{-47.58\lambda}.$$

We need to find the value of  $\lambda$ , within range of allowable parameters  $\Theta = (0, \infty)$ , which maximises  $L(\lambda; \mathbf{x})$ .



Differentiating,

$$\begin{aligned}\frac{dL}{d\lambda} &= 15\lambda^{14}e^{-47.58\lambda} - 47.58\lambda^{15}e^{-47.58\lambda} \\ &= \lambda^{14}(15 - 47.58\lambda)e^{-47.58\lambda}\end{aligned}$$

which is zero when  $\lambda = 0$  or  $\lambda = 15/47.58 \approx 0.32$ .

Differentiating,

$$\begin{aligned}\frac{dL}{d\lambda} &= 15\lambda^{14}e^{-47.58\lambda} - 47.58\lambda^{15}e^{-47.58\lambda} \\ &= \lambda^{14}(15 - 47.58\lambda)e^{-47.58\lambda}\end{aligned}$$

which is zero when  $\lambda = 0$  or  $\lambda = 15/47.58 \approx 0.32$ .

$\lambda = 0$  is outside of the allowed range  $\Theta = (0, \infty)$  of possible parameter values.

So, the only turning point of interest is  $\lambda = 15/47.58$ .

Differentiating again (with the details left to you), we end up with

$$\begin{aligned}\frac{d^2L}{d\lambda^2} &= (210\lambda^{13} - 1427.4\lambda^{14} + 2263.86\lambda^{15})e^{-47.58\lambda} \\ &= \lambda^{13} (210 - 1427.4\lambda + 2263.86\lambda^2) e^{-47.58\lambda}\end{aligned}$$

Evaluating at our turning point gives

$$\left. \frac{d^2L}{d\lambda^2} \right|_{\lambda=15/47.58} = \left( \frac{15}{47.58} \right)^{13} (-14.9996) e^{-15} < 0.$$

So, our turning point is a local maximum.

We have only one only turning point  $\lambda = 15/47.58$  within the allowable range. It is a local max, so it is also the global max.

Hence, the MLE for  $\lambda$ , given our sample data  $\mathbf{x}$ , is

$$\hat{\lambda} = 15/47.58 \approx 0.32.$$

# Mutations in DNA/RNA

The number of 'mutations per unit length' of DNA is modelled using a Poisson distribution, with an unknown parameter  $\theta \in (0, \infty)$ . The numbers of mutations found in disjoint sections of DNA are independent.

Using this model, find the likelihood function for the number of mutations present in a sample of  $n$  (disjoint) strands of DNA, each of which has unit length.

# Mutations in DNA/RNA

The number of 'mutations per unit length' of DNA is modelled using a Poisson distribution, with an unknown parameter  $\theta \in (0, \infty)$ . The numbers of mutations found in disjoint sections of DNA are independent.

Using this model, find the likelihood function for the number of mutations present in a sample of  $n$  (disjoint) strands of DNA, each of which has unit length.

Let  $X_i$  be the number of mutations in the  $i^{th}$  strand of DNA. So,

$$f_{X_i}(x_i; \theta) = \frac{e^{-\theta} \theta^{x_i}}{(x_i)!}$$

for  $x_i \in \mathbb{N} \cup \{0\}$ , and  $f_{X_i}(x_i) = 0$  if  $x_i \notin \mathbb{N} \cup \{0\}$ .

Since we assume the  $(X_i)$  are independent, the joint distribution of  $\mathbf{X} = (X_1, X_2, \dots, X_n)$  has probability function

$$\begin{aligned} f_{\mathbf{X}}(\mathbf{x}) &= \prod_{i=1}^n \frac{e^{-\theta} \theta^{x_i}}{(x_i)!} \\ &= \frac{1}{(x_1)!(x_2)! \dots (x_n)!} e^{-n\theta} \theta^{\sum_1^n x_i} \end{aligned}$$

provided all  $x_i \in \mathbb{N} \cup \{0\}$ , and zero otherwise. Therefore, our likelihood function is

$$L(\theta; \mathbf{x}) = \frac{1}{(x_1)!(x_2)! \dots (x_n)!} e^{-n\theta} \theta^{\sum_1^n x_i}.$$

The range of possible values for  $\theta$  is  $\Theta = (0, \infty)$ .

Suppose that at least one of the  $x_i$  is non-zero. Find the corresponding log-likelihood function, and hence find the maximum likelihood estimator of  $\theta$ .



Suppose that at least one of the  $x_i$  is non-zero. Find the corresponding log-likelihood function, and hence find the maximum likelihood estimator of  $\theta$ .

The log-likelihood function is  $\ell(\theta; \mathbf{x}) = \log L(\theta; \mathbf{x})$ , so

$$\begin{aligned}\log L(\theta, \mathbf{x}) &= \log \left( \frac{1}{(x_1)!(x_2)! \dots (x_n)!} e^{-n\theta} \theta^{\sum_{i=1}^n x_i} \right) \\ &= \sum_{i=1}^n (-\log(x_i)!) - n\theta + (\log \theta) \sum_{i=1}^n x_i.\end{aligned}$$

We now look to maximise  $\ell(\theta; \mathbf{x})$ , over  $\theta \in (0, \infty)$ .

Differentiating, we obtain

$$\frac{d\ell}{d\theta} = -n + \frac{1}{\theta} \sum_{i=1}^n x_i.$$

So, the only turning point of  $\ell(\theta, \mathbf{x})$  is at  $\theta = \frac{1}{n} \sum_{i=1}^n x_i$ .

Differentiating, we obtain

$$\frac{d\ell}{d\theta} = -n + \frac{1}{\theta} \sum_{i=1}^n x_i.$$

So, the only turning point of  $\ell(\theta, \mathbf{x})$  is at  $\theta = \frac{1}{n} \sum_{i=1}^n x_i$ .

Differentiating again, we have

$$\frac{d^2\ell}{d\theta^2} = -\frac{1}{\theta^2} \sum_{i=1}^n x_i.$$

Since the  $x_i$  count mutations we have  $x_i \geq 0$ , and since at least one  $x_i$  is non-zero we have  $\frac{d^2\ell}{d\theta^2} < 0$  (for all  $\theta$ ).

Therefore, the MLE of  $\theta$  given the data  $\mathbf{x}$  is  $\hat{\theta} = \frac{1}{n} \sum_{i=1}^n x_i$ .

Mutations rates were measured, for 11 HIV patients, and there were found to be

$$\mathbf{x} = \{19, 16, 37, 28, 24, 34, 37, 126, 32, 48, 45\}$$

mutations per  $10^4$  possible locations (i.e. 'per unit length'). Assuming the model above, calculate the maximum likelihood estimator of the mutation rate of HIV.

Mutations rates were measured, for 11 HIV patients, and there were found to be

$$\mathbf{x} = \{19, 16, 37, 28, 24, 34, 37, 126, 32, 48, 45\}$$

mutations per  $10^4$  possible locations (i.e. 'per unit length'). Assuming the model above, calculate the maximum likelihood estimator of the mutation rate of HIV.

The data has

$$\bar{x} = \frac{1}{11} \sum_{i=1}^{11} x_i = \frac{446}{11} \approx 41$$

so we conclude that the maximum likelihood estimator of the mutation rate  $\theta$  'per unit length', given this data, is  $\hat{\theta} = \frac{446}{11}$ .

# Mass spectroscopy

A sample of 15 amino acid molecules, which are all known to be of the same type (and therefore, the same mass), were reported to have masses

$$\mathbf{x} = \{63.7, 5.8, 112.0, 6.5, 73.0, 115.7, 73.0, 39.4, 98.8, 83.0, 198.5, 93.3, 132.3, 118.0, 63.3\}.$$

It is known that these molecules are either Alanine, which has mass 71.0, or Leucine, which has mass 113.1.

Given a molecule of mass  $\theta$ , the spectrometer is known to report its mass as  $X \sim N(\theta, 35^2)$ , independently for each molecule.

Using this model, and the data above, find the likelihoods of Alanine and Leucine. Specify which of these maximises the likelihood.

Our model, for the reported mass  $X$  of a single molecule with (real) weight  $\theta$ , is  $X \sim N(\theta, 35^2)$  and the p.d.f. of a single data point is

$$f_{X_i}(x_i) = \frac{1}{(2\pi)^{1/2}35} \exp\left(-\frac{(x_i - \theta)^2}{2 \times 35^2}\right).$$

Therefore, the p.d.f. of the reported masses  $\mathbf{X} = (X_1, \dots, X_n)$  of  $n$  molecules is

$$f_{\mathbf{X}}(\mathbf{x}) = \prod_{i=1}^n f_{X_i}(x_i) = \frac{1}{(2\pi)^{n/2}35^n} \exp\left(-\frac{1}{2450} \sum_{i=1}^n (x_i - \theta)^2\right).$$

We know that, in reality,  $\theta$  must be one of only two different values; 71.0 (for Alanine) and 113.1 (for Leucine). Therefore, our likelihood function is

$$L(\theta; \mathbf{x}) = \frac{1}{(2\pi)^{n/2} 35^n} \exp \left( -\frac{1}{2450} \sum_{i=1}^n (x_i - \theta)^2 \right)$$

and the possible range of values for  $\theta$  is the two point set  $\Theta = \{71.0, 113.1\}$ . We need to find out which of these two values maximises the likelihood.



Our data  $\mathbf{x}$  contains  $n = 15$  data points. A short calculation (use e.g. R) shows that

$$\frac{1}{2450} \sum_{i=1}^{15} (x_i - 71.0)^2 \approx 12.70, \quad \frac{1}{2450} \sum_{i=1}^{15} (x_i - 113.1)^2 \approx 20.41.$$

and, therefore, that

$$L(71.0; \mathbf{x}) \approx 2.19 \times 10^{-34}, \quad L(113.1; \mathbf{x}) = 9.90 \times 10^{-38}.$$

We conclude that  $\theta = 71.0$  has (much) greater likelihood than  $\theta = 113.1$ , so we expect that the molecules sampled are Alanine.

# Rainfall

Find the maximum likelihood estimator of the parameter vector  $\theta = (\mu, \sigma^2)$  when the data  $\mathbf{x} = (x_1, x_2, \dots, x_n)$  are modelled as i.i.d. samples from a normal distribution  $N(\mu, \sigma^2)$ .

Our parameter vector is  $\theta = (\mu, \sigma^2)$ , so let us write  $v = \sigma^2$  to avoid confusion. This means we are interested in the parameters

$$\theta = (\mu, v),$$

and the range of possible values of  $\theta$  is  $\Theta = \mathbb{R} \times (0, \infty)$ .

The p.d.f. of the univariate normal distribution  $N(\mu, \nu)$  is

$$f_X(x) = \frac{1}{\sqrt{2\pi\nu}} e^{-(x-\mu)^2/2\nu}.$$

Writing  $\mathbf{X} = (X_1, \dots, X_n)$ , where the  $X_i$  are i.i.d. univariate  $N(\mu, \nu)$  random variables, the likelihood function of  $\mathbf{X}$  is

$$L(\boldsymbol{\theta}; \mathbf{x}) = f_{\mathbf{X}}(\mathbf{x}) = \frac{1}{(2\pi\nu)^{n/2}} \exp\left(-\frac{1}{2\nu} \sum_{i=1}^n (x_i - \mu)^2\right).$$

Therefore, the log likelihood is

$$\ell(\boldsymbol{\theta}; \mathbf{x}) = -\frac{n}{2} (\log(2\pi) + \log(\nu)) - \frac{1}{2\nu} \sum_{i=1}^n (x_i - \mu)^2.$$

We now look to maximise  $\ell(\theta; \mathbf{x})$  over  $\theta \in \Theta$ . The partial derivatives are

$$\begin{aligned}\frac{\partial \ell}{\partial \mu} &= \frac{1}{v} \sum_{i=1}^n (x_i - \mu) = \frac{1}{v} \left( \sum_{i=1}^n x_i - n\mu \right) \\ \frac{\partial \ell}{\partial v} &= -\frac{n}{2v} + \frac{1}{2v^2} \sum_{i=1}^n (x_i - \mu)^2.\end{aligned}$$

Solving  $\frac{\partial \ell}{\partial \mu} = 0$  gives  $\mu = \frac{1}{n} \sum_{i=1}^n x_i = \bar{x}$ . Solving  $\frac{\partial \ell}{\partial v} = 0$  gives  $v = \frac{1}{n} \sum_{i=1}^n (x_i - \mu)^2$ . So both partial derivatives will be zero if and only if

$$\mu = \bar{x}, \quad v = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2. \quad (1)$$

This gives us the value of  $\theta = (\mu, v)$  at the (single) turning point of  $\ell$ .

Next, we use the Hessian matrix to check if this point is a local maximum. We have

$$\begin{aligned}\frac{\partial^2 \ell}{\partial \mu^2} &= -\frac{n}{v} \\ \frac{\partial^2 \ell}{\partial \mu \partial v} &= \frac{-1}{v^2} \left( \sum_{i=1}^n x_i - n\mu \right) \\ \frac{\partial^2 \ell}{\partial v^2} &= \frac{n}{2v^2} - \frac{1}{v^3} \sum_{i=1}^n (x_i - \mu)^2\end{aligned}$$

Evaluating these at our turning point, we get

$$\left. \frac{\partial^2 \ell}{\partial \mu^2} \right|_{(1)} = -\frac{n}{\hat{v}}$$

$$\left. \frac{\partial^2 \ell}{\partial \mu \partial v} \right|_{(1)} = \frac{-1}{v^2} \left( \sum_{i=1}^n x_i - n\bar{x} \right) = 0$$

$$\left. \frac{\partial^2 \ell}{\partial v^2} \right|_{(1)} = \frac{n}{2v^2} - \frac{1}{v^3} \sum_{i=1}^n (x_i - \bar{x})^2 = \frac{n}{2v^2} - \frac{1}{v^3} n\hat{v} = \frac{-n}{2v^2}$$

so

$$H = \begin{pmatrix} -\frac{n}{v} & 0 \\ 0 & \frac{-n}{2v^2} \end{pmatrix}.$$

Since  $-\frac{n}{v} < 0$  and  $\det H = \frac{n^2}{2v^3} > 0$ , our turning point (1) is a local maximum. Since it is the only turning point, it is also the global maximum. Hence, the MLE is

$$\hat{\mu} = \bar{x}$$
$$\hat{\sigma}^2 = \hat{v} = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2.$$

Note  $\hat{\mu}$  is the sample mean, and  $\hat{\sigma}^2$  is the (biased) sample variance.

For the years 1985-2015, the amount of rainfall (in millimeters) recorded as falling on Sheffield in December is as follows:

{78.0, 142.3, 38.2, 36.0, 159.1, 136.0, 78.4, 67.4, 171.4, 103.9, 70.4, 98.2, 79.4, 57.9, 135.6, 118.0, 28.0, 129.8, 106.5, 46.3, 56.7, 114.0, 74.9, 52.8, 66.1, 18.8, 124.6, 136.0, 69.8, 102.0, 121.2}

Assuming that we choose to model the amount of rainfall in Sheffield each December by a normal distribution, find the maximum likelihood estimators for  $\mu$  and  $\sigma^2$ .



For the years 1985-2015, the amount of rainfall (in milimeters) recorded as falling on Sheffield in December is as follows:

{78.0, 142.3, 38.2, 36.0, 159.1, 136.0, 78.4, 67.4, 171.4, 103.9, 70.4, 98.2, 79.4, 57.9, 135.6, 118.0, 28.0, 129.8, 106.5, 46.3, 56.7, 114.0, 74.9, 52.8, 66.1, 18.8, 124.6, 136.0, 69.8, 102.0, 121.2}

Assuming that we choose to model the amount of rainfall in Sheffield each December by a normal distribution, find the maximum likelihood estimators for  $\mu$  and  $\sigma^2$ .

The data has  $n = 30$ , and

$$\bar{x} = \frac{1}{30} \sum_{i=1}^{30} x_i \approx 93.9, \quad \frac{1}{30} \sum_{i=1}^{30} (x_i - \bar{x})^2 \approx 1631.2 \approx 40.4^2$$

So we conclude that, according to our model, the maximum likelihood estimators are  $\hat{\mu} \approx 93.9$  and  $\hat{\sigma}^2 \approx 40.4^2$ , which means that Sheffield receives a  $N(93.9, 40.4^2)$  quantity of rainfall, in millimetres, each December.